

# **RedundancyMiner**

## **User's Manual**

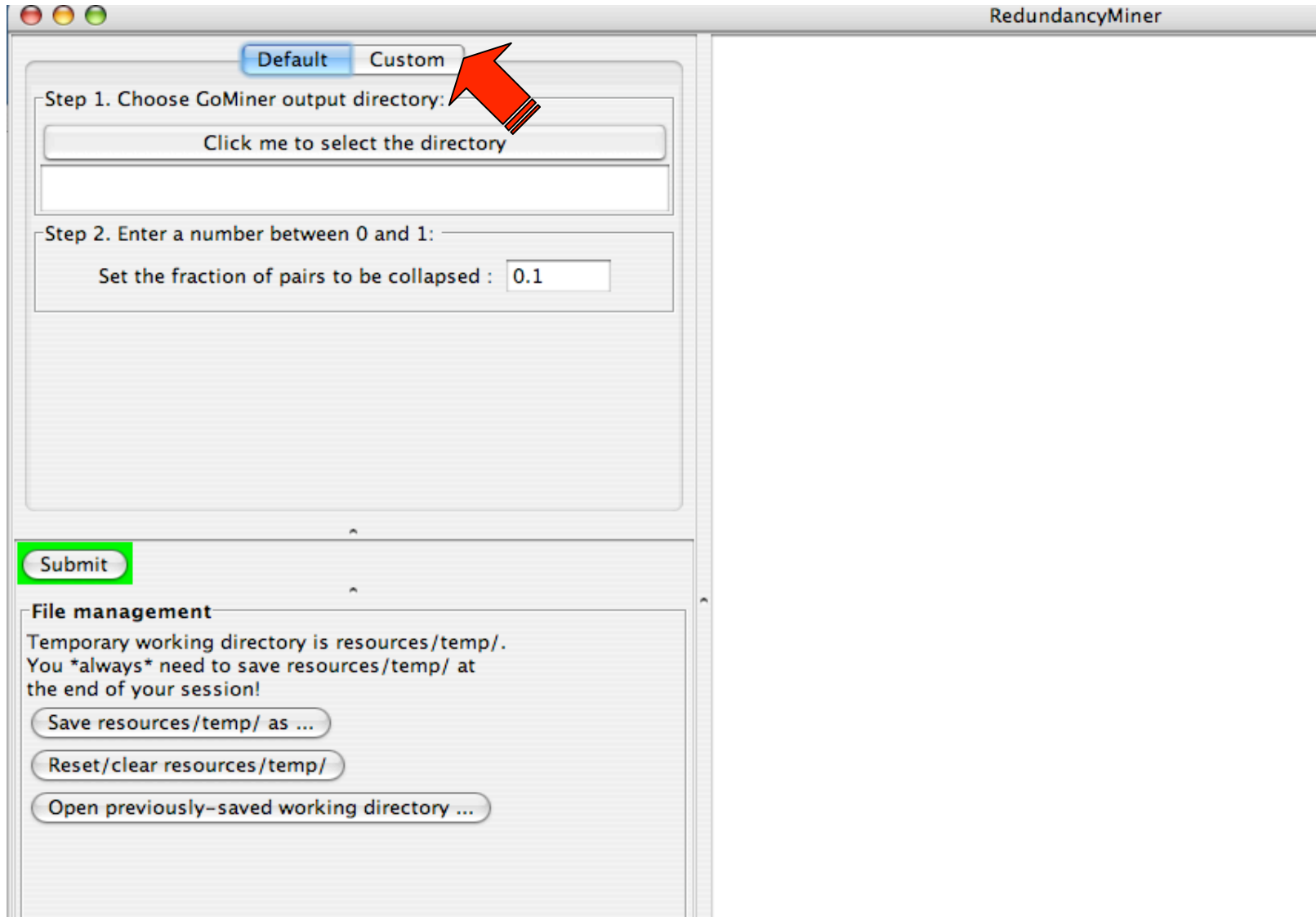
To follow along with the presentation, please download a simplified version of the HTGM result directory that is exemplified in the user's manual.

This is available from  
<http://discover.nci.nih.gov/rm/supplementaryMaterials.html>

# Custom Mode

- Computing the similarity matrix
  - Select .CIM file
  - Select .gce or .tvf file
  - Submit the job
- Clustering
  - Select P-value threshold
  - Submit the job
- File management
  - Save the working directory
  - Retrieve the working directory

# Custom Mode: Select Custom Mode



The screenshot shows the RedundancyMiner application window. At the top, there are two tabs: "Default" and "Custom". A red arrow points to the "Custom" tab. Below the tabs, the interface is divided into two main sections. The top section contains two steps: "Step 1. Choose GoMiner output directory:" with a button labeled "Click me to select the directory" and an empty text box below it; and "Step 2. Enter a number between 0 and 1:" with a label "Set the fraction of pairs to be collapsed :" and a text box containing "0.1". The bottom section is titled "File management" and contains a warning message: "Temporary working directory is resources/temp/. You \*always\* need to save resources/temp/ at the end of your session!". Below the message are three buttons: "Save resources/temp/ as ...", "Reset/clear resources/temp/", and "Open previously-saved working directory ...". A green box highlights the "Submit" button located between the two main sections.

RedundancyMiner

Default Custom

Step 1. Choose GoMiner output directory:

Click me to select the directory

Step 2. Enter a number between 0 and 1:

Set the fraction of pairs to be collapsed : 0.1

Submit

**File management**

Temporary working directory is resources/temp/.  
You \*always\* need to save resources/temp/ at  
the end of your session!

Save resources/temp/ as ...

Reset/clear resources/temp/

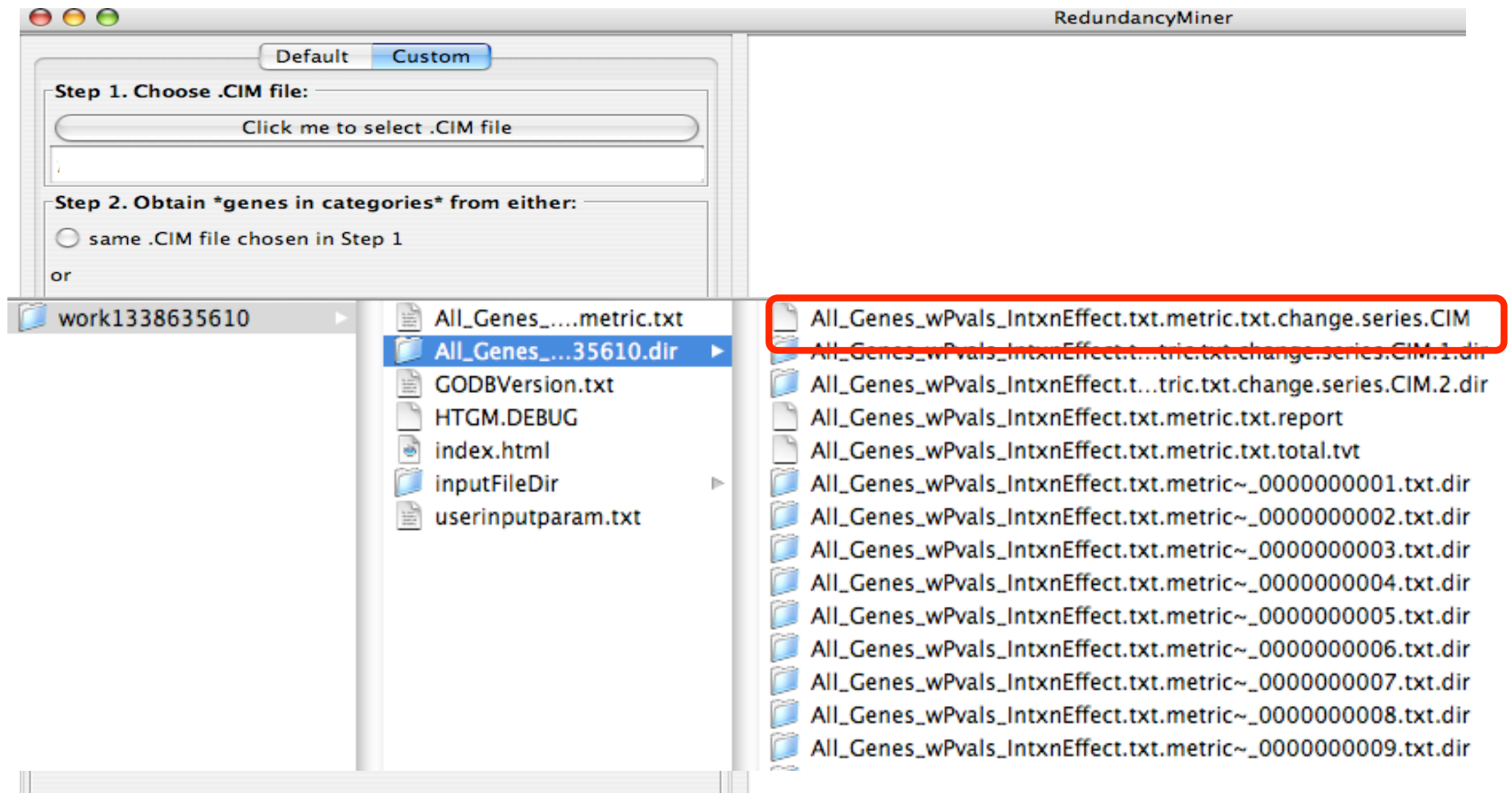
Open previously-saved working directory ...

# Compute the similarity matrix: Choose .CIM (Step 1)

The screenshot shows the RedundancyMiner application window. The title bar reads "RedundancyMiner". Inside the window, there are two tabs: "Default" and "Custom", with "Custom" being the active tab. The interface is divided into several sections:

- Step 1. Choose .CIM file:** This section contains a button labeled "Click me to select .CIM file". A red arrow points to this button.
- Step 2. Obtain \*genes in categories\* from either:** This section contains two radio buttons:
  - ☐ same .CIM file chosen in Step 1
  - ☒ .gce or .tvt fileBelow the radio buttons is a button labeled "Click me to select .gce or .tvt file".
- The selected \*genes in categories\* file is:** This section contains an empty text box.
- Submit:** A button with a green border.
- File management:** This section contains the following text and buttons:
  - Temporary working directory is resources/temp/.
  - You \*always\* need to save resources/temp/ at the end of your session!
  - Save resources/temp/ as ...
  - Reset/clear resources/temp/
  - Open previously-saved working directory ...

# Compute the similarity matrix: Choose .CIM (Step 2)



# Compute the similarity matrix: Choose .gce or .tvf (Step 1)

The screenshot shows the RedundancyMiner web interface. At the top, there are two tabs: "Default" and "Custom". Below the tabs, the interface is divided into two main sections. The left section contains the steps for computing the similarity matrix. Step 1 is "Choose .CIM file:", followed by a button "Click me to select .CIM file". Below this button, a text box contains the file path ".dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.change.series.CIM", which is highlighted with a red rectangle. Step 2 is "Obtain \*genes in categories\* from either:", followed by two radio buttons. The first radio button is "same .CIM file chosen in Step 1". The second radio button is ".gce or .tvf file", which is selected. Below the radio buttons, there is a button "Click me to select .gce or .tvf file". Below this button, there is a text box "The selected \*genes in categories\* file is:" with an empty input field. A red arrow points to this input field. At the bottom of the left section, there is a green "Submit" button. The right section of the interface is titled "File management" and contains the text "Temporary working directory is resources/temp/. You \*always\* need to save resources/temp/ at the end of your session!". Below this text, there are three buttons: "Save resources/temp/ as ...", "Reset/clear resources/temp/", and "Open previously-saved working directory ...".

RedundancyMiner

Default Custom

Step 1. Choose .CIM file:

Click me to select .CIM file

.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.change.series.CIM

Step 2. Obtain \*genes in categories\* from either:

☐ same .CIM file chosen in Step 1

or

☒ .gce or .tvf file

Click me to select .gce or .tvf file

The selected \*genes in categories\* file is:

Submit

File management

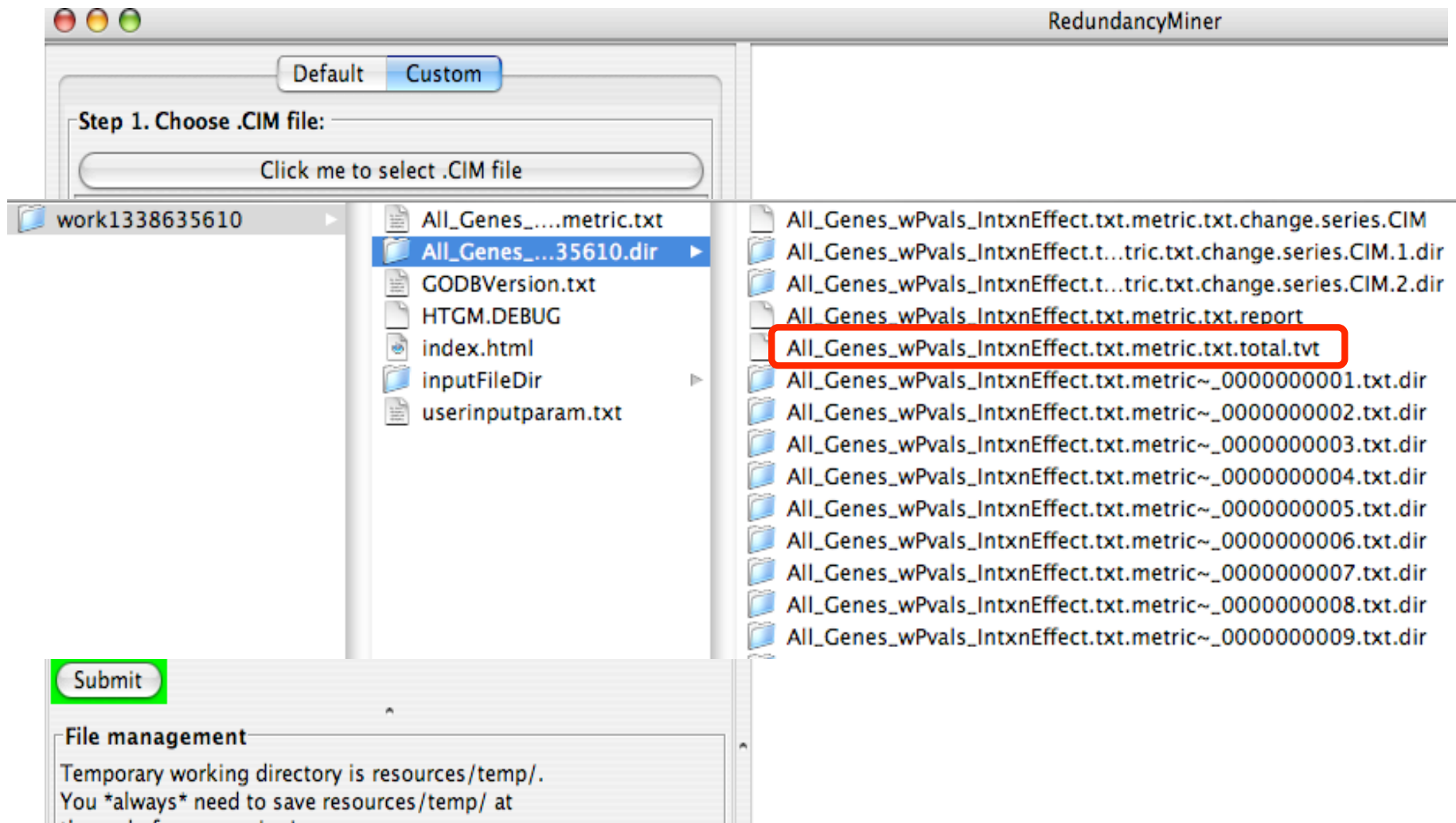
Temporary working directory is resources/temp/.  
You \*always\* need to save resources/temp/ at the end of your session!

Save resources/temp/ as ...

Reset/clear resources/temp/

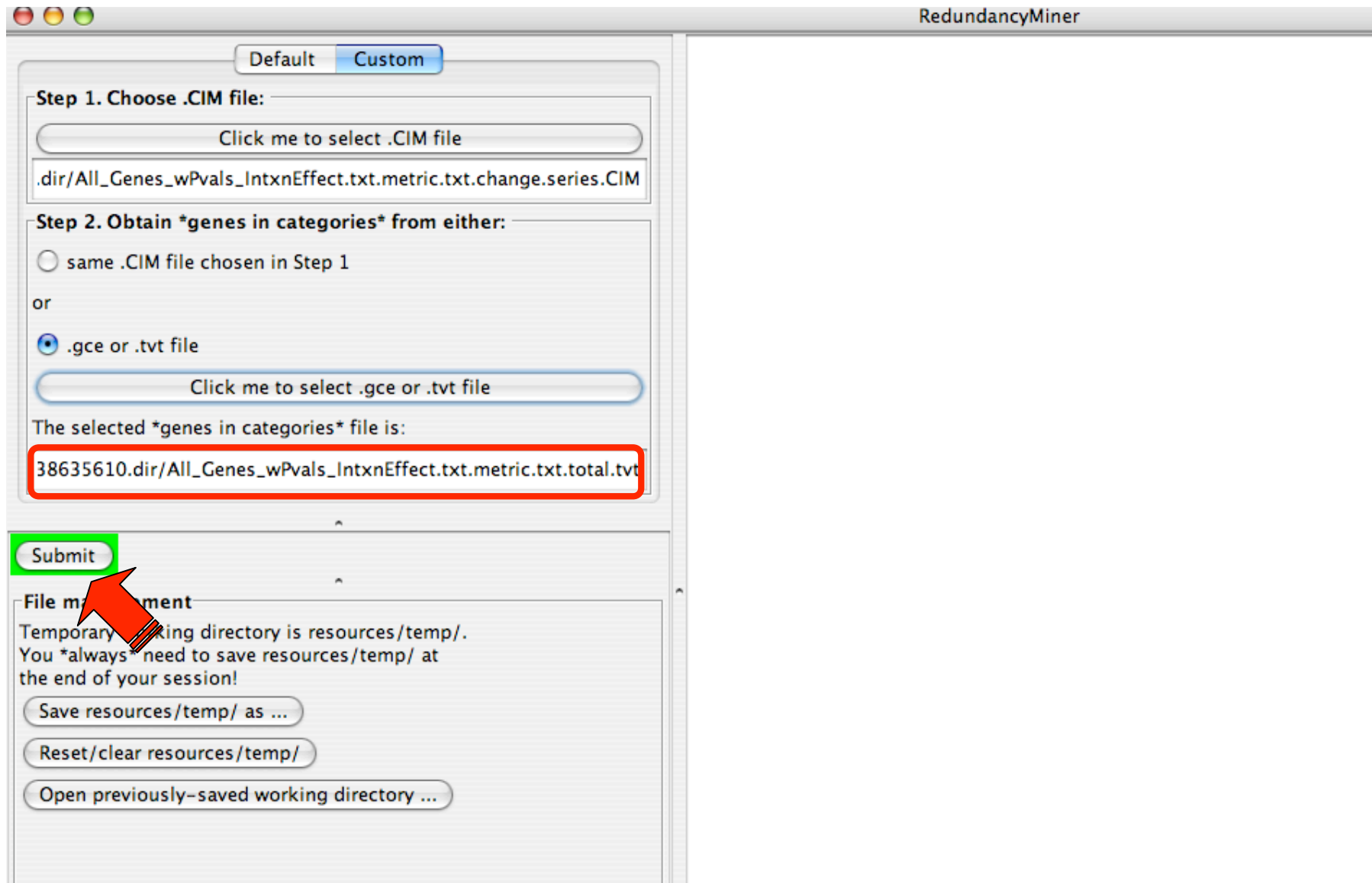
Open previously-saved working directory ...

Compute the similarity matrix: Choose .gce or tvf (Step 2)





# Compute the similarity matrix: Submit the job



The screenshot shows the RedundancyMiner application window. The 'Custom' tab is selected. Step 1 shows a .CIM file path. Step 2 shows the selection of a .gce or .txt file. The 'Submit' button is highlighted with a green box and a red arrow. The file path for the selected file is highlighted with a red box.

RedundancyMiner

Default Custom

Step 1. Choose .CIM file:

Click me to select .CIM file

.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.change.series.CIM

Step 2. Obtain \*genes in categories\* from either:

☐ same .CIM file chosen in Step 1

or

☒ .gce or .txt file

Click me to select .gce or .txt file

The selected \*genes in categories\* file is:

38635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.total.txt

Submit

File management

Temporary working directory is resources/temp/.  
You \*always\* need to save resources/temp/ at the end of your session!

Save resources/temp/ as ...

Reset/clear resources/temp/

Open previously-saved working directory ...

## Cluster: Select P-value threshold (Step 1)

A glitch in the java interface keeps the Threshold box from appearing automatically.

**If you click on the interface in this area, it will appear.**

**Threshold box from appearing automatically.**

**If you click on the interface in this area, it will appear.**

**Less collapsing**

**More collapsing**

**File management**  
 Temporary working directory is resources/temp/.  
 You \*always\* need to save resources/temp/ at the end of your session!

Save resources/temp/ as ...  
 Reset/clear resources/temp/  
 Open previously-saved working directory ...

**Step 1. Choose .CIM file:**  
 Click me to select .CIM file  
 .dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.change.series.CIM

**Step 2. Obtain \*genes in categories\* from either:**  
☐ same .CIM file chosen in Step 1  
 or  
☒ .gce or .tvt file  
 Click me to select .gce or .tvt file  
 The selected \*genes in categories\* file is:  
 38635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.total.tvt

**Step 3. Set log10(P-Value) Threshold:**  
 log10(P-Value) Threshold:

**Submit**

Reading data ...  
 Calculating similarities ...  
 Generating data for FDR calculation if needed ...  
 The process finishes at Tue Dec 23 15:15:55 2008  
 Temporary directory is resources/temp/ ...  
 Obtaining statistics ...  
 The process of obtaining statistics starts at Tue Dec 23 15:15:55 2008  
 Total number of pairs is 16653  
 Only show those with P-value less than 0.5

Rank	log10(P-Value)
0	-9.223372e+16
1	-9.223372e+16
3	-9.223372e+16
7	-9.223372e+16
15	-9.223372e+16
31	-9.223372e+16
63	-9.223372e+16
127	-2.254500e+02
255	-1.109700e+02
511	-3.832000e+01
1023	-1.380000e+01
2047	-3.950000e+00
4095	-1.130000e+00

Rank	log10(P-value) Thresh	Number of Collapsed Pairs
0	-9.223372e+16	90
1	-9.223372e+16	90
3	-9.223372e+16	90
7	-9.223372e+16	90
15	-9.223372e+16	90
31	-9.223372e+16	90
63	-9.223372e+16	90
127	-2.254500e+02	128
255	-1.109700e+02	256
511	-3.832000e+01	512
1023	-1.380000e+01	1023
2047	-3.950000e+00	2047
4095	-1.130000e+00	4095
8191	-8.000000e-02	8255
16383	0.000000e+00	16653

The process of obtaining statistics finishes at Tue Dec 23 15:15:58 2008

Cluster: Select P-value threshold (Step 2)

**A glitch in the java interface keeps the Threshold box from appearing automatically.**

**Threshold box from appearing**

**automatically.**

**If you click on the interface in this area, it will appear.**

Default Custom

**Step 1. Choose .CIM file:**

Click me to select .CIM file

.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.change.series.CIM

**Step 2. Obtain \*genes in categories\* from either:**

☐ same .CIM file chosen in Step 1

or

☒ .gce or .txt file

Click me to select .gce or .txt file

The selected \*genes in categories\* file is:

38635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.total.txt

**Step 3. Set log10(P-Value) Threshold:**

log10(P-Value) Threshold:

Submit

**File management**

Temporary working directory is resources/temp/.  
You \*always\* need to save resources/temp/ at the end of your session!

Save resources/temp/ as ...

Reset/clear resources/temp/

Open previously-saved working directory ...

Reading data ....  
Saving hit genes  
resources/temp/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.total.txt.All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.total.txt  
Generating data for FDR calculation if needed ....  
The process finishes at Tue Dec 23 15:15:55 2008  
Temporary directory is resources/temp/ ...  
Obtaining statistics  
The process of obtaining statistics starts at Tue Dec 23 15:15:58 2008  
Total number of pairs is 16653  
Only show those with P-value less than 0.5

Rank	log10(P-Value)
0	-9.223372e+16
1	-9.223372e+16
3	-9.223372e+16
7	-9.223372e+16
15	-9.223372e+16
31	-9.223372e+16
63	-9.223372e+16
127	-2.254500e+02
255	-1.109700e+02
511	-3.832000e+01
1023	-1.380000e+01
2047	-3.950000e+00
4095	-1.130000e+00

Rank	log10(P-value) Threshold	Number of Collapsed Pairs
0	-9.223372e+16	90
1	-9.223372e+16	90
3	-9.223372e+16	0
7	-9.223372e+16	0
15	-9.223372e+16	0
31	-9.223372e+16	0
63	-9.223372e+16	0
127	-2.254500e+02	28
255	-1.109700e+02	256
511	-3.832000e+01	512
1023	-1.380000e+01	1024
2047	-3.950000e+00	2050
4095	-1.130000e+00	4104
8191	-8.000000e-02	8255
16383	0.000000e+00	16653

Copy

The process of obtaining statistics finishes at Tue Dec 23 15:15:58 2008

# Cluster: Select P-value threshold (Step 3)

**RedundancyMiner**

Default Custom

**Step 1. Choose .CIM file:**

Click me to select .CIM file

.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.change.series.CIM

**Step 2. Obtain \*genes in categories\* from either:**

☐ same .CIM file chosen in Step 1

or

☒ .gce or .txt file

Click me to select .gce or .txt file

The selected \*genes in categories\* file is:

38635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.total.txt

**Step 3. Set log10(P-Value) Threshold:**

log10(P-Value) Threshold:

Paste

**Submit**

**File management**

Temporary working directory is resources/temp/.  
You \*always\* need to save resources/temp/ at the end of your session!

Save resources/temp/ as ...

Reset/clear resources/temp/

Open previously-saved working directory ...

Reading data ....  
Calculating similarities ....  
resources/temp/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.total.txt.All\_Gen  
put.s2s  
Generating data for FDR calculation if needed ....  
The process finishes at Tue Dec 23 15:15:55 2008  
Temporary directory is resources/temp/ ...  
Obtaining statistics  
The process of obtaining statistics starts at Tue Dec 23 15:15:58 2008

Total number of pairs is 16653  
Only show those with P-value less than 0.5

Rank	log10(P-Value)
0	-9.223372e+16
1	-9.223372e+16
3	-9.223372e+16
7	-9.223372e+16
15	-9.223372e+16
31	-9.223372e+16
63	-9.223372e+16
127	-2.254500e+02
255	-1.109700e+02
511	-3.832000e+01
1023	-1.380000e+01
2047	-3.950000e+00
4095	-1.130000e+00

Rank	log10(P-value) Threshold	Number of Collapse
0	-9.223372e+16	90
1	-9.223372e+16	90
3	-9.223372e+16	90
7	-9.223372e+16	90
15	-9.223372e+16	90
31	-9.223372e+16	90
63	-9.223372e+16	90
127	-2.254500e+02	128
255	-1.109700e+02	256
511	-3.832000e+01	512
1023	-1.380000e+01	1024
2047	-3.950000e+00	2050
4095	-1.130000e+00	4104
8191	-8.000000e-02	8255
16383	0.000000e+00	16653

The process of obtaining statistics finishes at Tue Dec 23 15:15:58 2008

# Cluster: Submit the job

The screenshot shows the RedundancyMiner application window. The interface is divided into several sections:

- Default/Custom:** A tab at the top with 'Custom' selected.
- Step 1. Choose .CIM file:** A text box containing the file path: `.dir/All_Genes_wPvals_IntxnEffect.txt.metric.txt.change.series.CIM`.
- Step 2. Obtain \*genes in categories\* from either:** Two radio buttons are present. The first is 'same .CIM file chosen in Step 1'. The second is '.gce or .txt file', which is selected. Below it is a text box containing the file path: `38635610.dir/All_Genes_wPvals_IntxnEffect.txt.metric.txt.total.txt`.
- Step 3. Set log10(P-Value) Threshold:** A text box containing the value: `-9.223372e+16`.
- Submit:** A green button with a red arrow pointing to it.
- File management:** A section at the bottom with three buttons: 'Save resources/temp/ as ...', 'Reset/clear resources/temp/', and 'Open previously-saved working directory ...'.
- Output Log:** A text area on the right side of the window showing the progress of the job.

**Output Log Content:**

```
Reading data ....
Calculating similarities ....
resources/temp/All_Genes_wPvals_IntxnEffect.txt.metric.txt.total.txt.All_Gen
put.s2s
Generating data for FDR calculation if needed ....
The process finishes at Tue Dec 23 15:15:55 2008
Temporary directory is resources/temp/ ...
Obtaining statistics
The process of obtaining statistics starts at Tue Dec 23 15:15:58 2008

Total number of pairs is 16653
Only show those with P-value less than 0.5
```

Rank	log10(P-Value)
0	-9.223372e+16
1	-9.223372e+16
3	-9.223372e+16
7	-9.223372e+16
15	-9.223372e+16
31	-9.223372e+16
63	-9.223372e+16
127	-2.254500e+02
255	-1.109700e+02
511	-3.832000e+01
1023	-1.380000e+01
2047	-3.950000e+00
4095	-1.130000e+00

Rank	log10(P-value) Threshold	Number of Collapse
0	-9.223372e+16	90
1	-9.223372e+16	90
3	-9.223372e+16	90
7	-9.223372e+16	90
15	-9.223372e+16	90
31	-9.223372e+16	90
63	-9.223372e+16	90
127	-2.254500e+02	128
255	-1.109700e+02	256
511	-3.832000e+01	512
1023	-1.380000e+01	1024
2047	-3.950000e+00	2050
4095	-1.130000e+00	4104
8191	-8.000000e-02	8255
16383	0.000000e+00	16653

The process of obtaining statistics finishes at Tue Dec 23 15:15:58 2008



# Cluster: Completed

Default Custom

Step 1. Choose .CIM file:

Click me to select .CIM file

.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.change.series.CIM

Step 2. Obtain \*genes in categories\* from either:

☐ same .CIM file chosen in Step 1

or

☒ .gce or .txt file

Click me to select .gce or .txt file

The selected \*genes in categories\* file is:

38635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.total.txt

Step 3. Set log<sub>10</sub>(P-Value) Threshold:

log<sub>10</sub>(P-Value) Threshold: -9.223372e+16

Submit

File management

Temporary working directory is resources/temp/.  
You \*always\* need to save resources/temp/ at the end of your session!

Save resources/temp/ as ...

Reset/clear resources/temp/

Open previously-saved working directory ...

RedundancyMiner

0	-9.223372e+16	
1	-9.223372e+16	
3	-9.223372e+16	
7	-9.223372e+16	
15	-9.223372e+16	
31	-9.223372e+16	
63	-9.223372e+16	
127	-2.254500e+02	
255	-1.109700e+02	
511	-3.832000e+01	
1023	-1.380000e+01	
2047	-3.950000e+00	
4095	-1.130000e+00	

Rank	log <sub>10</sub> (P-value) Threshold	Number of Collapsed Pairs
0	-9.223372e+16	90
1	-9.223372e+16	90
3	-9.223372e+16	90
7	-9.223372e+16	90
15	-9.223372e+16	90
31	-9.223372e+16	90
63	-9.223372e+16	90
127	-2.254500e+02	128
255	-1.109700e+02	256
511	-3.832000e+01	512
1023	-1.380000e+01	1024
2047	-3.950000e+00	2050
4095	-1.130000e+00	4104
8191	-8.000000e-02	8255
16382	0.000000e+00	16552

The process of obtaining statistics finishes at Tue Dec 23 15:44:33 2008

Temporary directory is resources/temp/ ...

Multicustering

The process starts at Tue Dec 23 15:44:35 2008

Initializing clusters ....

Building clusters ....

Printing clusters ....

The process finishes at Tue Dec 23 15:44:36 2008

Collapsing CIM according to clustering results and storing in resources/temp/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.change.series.CIM.0.-9.223372E16.0.0.0.0

The process starts at Tue Dec 23 15:44:37 2008

The number of categories:

Before collapsing: 183

After collapsing: 162

Compression ratio: 1.13

The process finishes at Tue Dec 23 15:44:41 2008

# File management: Save the working directory (Step 1)

The screenshot displays the RedundancyMiner application window. The interface is divided into several sections:

- Default/Custom:** A toggle switch at the top left.
- Step 1. Choose .CIM file:** A section with a button "Click me to select .CIM file" and a text field containing ".dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.change.series.CIM".
- Step 2. Obtain \*genes in categories\* from either:** Two radio buttons: "same .CIM file chosen in Step 1" (unselected) and ".gce or .txt file" (selected). Below is a button "Click me to select .gce or .txt file" and a text field containing "38635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.total.txt".
- Step 3. Set log10(P-Value) Threshold:** A label "log10(P-Value) Threshold:" followed by a text field containing "-9.223372e+16".
- Submit:** A green button with the text "Submit".
- File management:** A section at the bottom left containing text: "Temporary working directory is resources/temp/. You \*always\* need to save resources/temp/ at the end of your session!". Below this text are three buttons: "Save resources/temp/ as ...", "Reset/clear resources/temp/", and "Open previously-saved working directory ...". A red arrow points to the "Save resources/temp/ as ..." button.
- Log/Output:** A large text area on the right side of the window. It contains a table of data, a summary of the process, and a detailed log of operations.

**Table Data:**

Rank	log10(P-value)	Threshold	Number of Collapsed Pairs
0	-9.223372e+16	90	
1	-9.223372e+16	90	
3	-9.223372e+16	90	
7	-9.223372e+16	90	
15	-9.223372e+16	90	
31	-9.223372e+16	90	
63	-9.223372e+16	90	
127	-2.254500e+02	128	
255	-1.109700e+02	256	
511	-3.832000e+01	512	
1023	-1.380000e+01	1024	
2047	-3.950000e+00	2050	
4095	-1.130000e+00	4104	
8191	-8.000000e-02	8255	
16383	0.000000e+00	16653	

**Log/Output Text:**

The process of obtaining statistics finishes at Tue Dec 23 15:44:33 2008

Temporary directory is resources/temp/ ...

Multiclustering

The process starts at Tue Dec 23 15:44:35 2008

Initializing clusters ....

Building clusters ....

Printing clusters ....

The process finishes at Tue Dec 23 15:44:36 2008

Collapsing CIM according to clustering results and storing in resources/temp/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.change.series.CIM.0.-9.223372E16.0.0.0.0 ...

The process starts at Tue Dec 23 15:44:37 2008

The number of categories:

Before collapsing: 183

After collapsing: 162

Compression ratio: 1.13

The process finishes at Tue Dec 23 15:44:41 2008

# File management: Save the working directory (Step 2)

RedundancyMiner

Default Custom

Step 1. Choose .CIM file:

Click me to select .CIM file

.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.change.series.CIM

Step 2. Obtain \*genes in categories\* from either:

☐ same .CIM file chosen in Step 1

or

☒ .gce or .txt file

Click me to select .gce or .txt file

The selected \*genes in categories\* file is:

38635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.total

Step 3. Set log10(P-Value) Threshold:

log10(P-Value) Threshold: -9.223372e+16

Submit

File management

Temporary working directory is resources/temp/.  
You \*always\* need to save resources/temp/ at the end of your session!

Save resources/temp/ as ...

Reset/clear resources/temp/

Open previously-saved working directory ...

Save

Save As: myRMresult

gomineroutput

Name	Date Modified
work1338635610	Friday, November 21, 2008 3:52 PM

File Format: All Files

New Folder Cancel Save

Initializing clusters ....  
Building clusters ....  
Printing clusters ....  
The process finishes at Tue Dec 23 15:44:36 2008  
Collapsing CIM according to clustering results and storing in  
resources/temp/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.change.series.CIM.0.-9.223372E16.0.0.0.0 ...  
The process starts at Tue Dec 23 15:44:37 2008  
The number of categories:  
Before collapsing: 183  
After collapsing: 162  
Compression ratio: 1.13

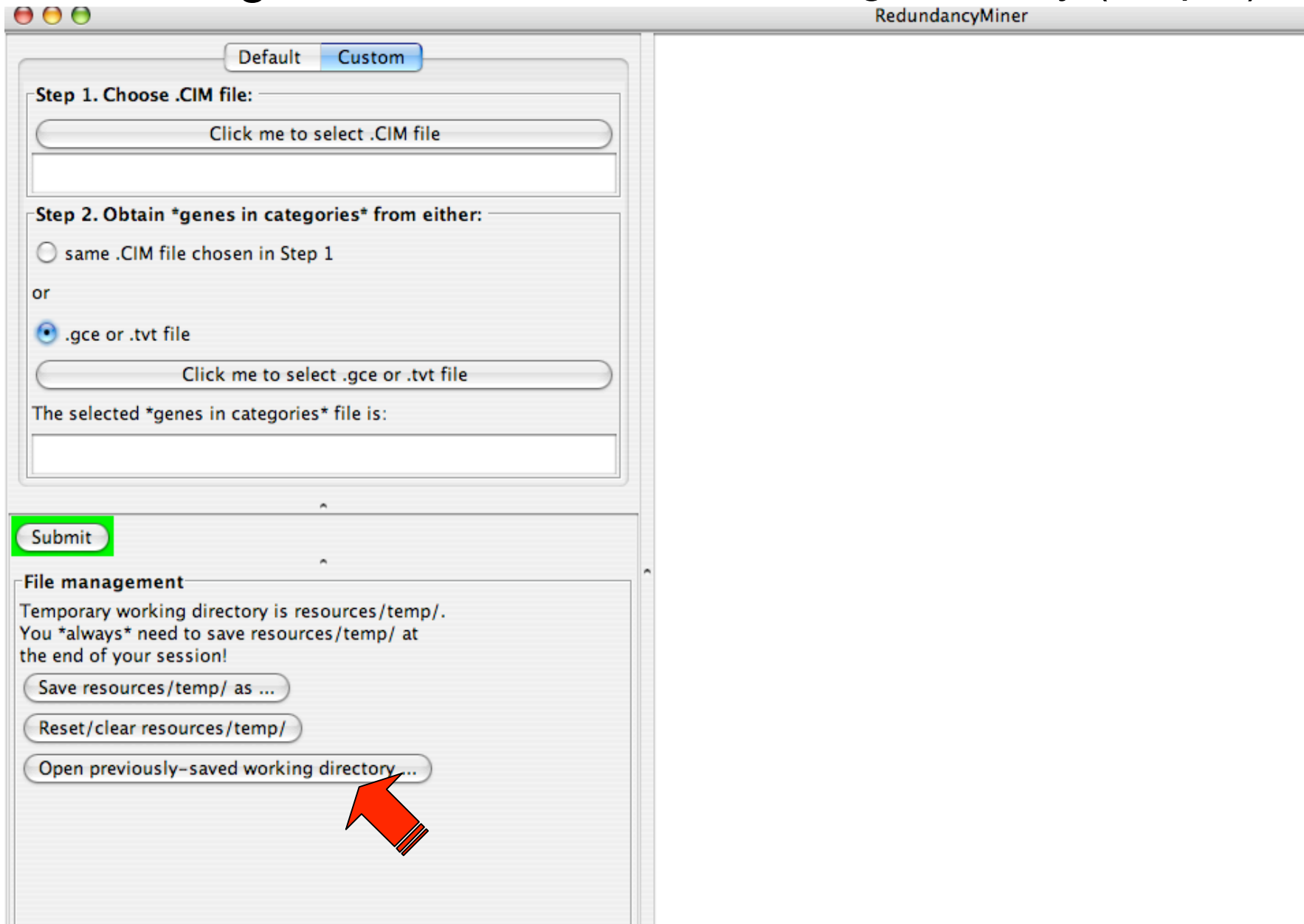


# File management: Save completed

The screenshot displays the RedundancyMiner application window. The interface is divided into several sections:

- Default Custom** tabs at the top.
- Step 1. Choose .CIM file:** A button labeled "Click me to select .CIM file" and a text field containing ".dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.change.series.CIM".
- Step 2. Obtain \*genes in categories\* from either:** Two radio buttons: "same .CIM file chosen in Step 1" (unselected) and ".gce or .txt file" (selected). Below is a button "Click me to select .gce or .txt file" and a text field showing "38635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.total.txt".
- Step 3. Set log<sub>10</sub>(P-Value) Threshold:** A text field labeled "log<sub>10</sub>(P-Value) Threshold:" with the value "-9.223372e+16".
- Submit** button: A green button with a white border.
- File management** section: A text box stating "Temporary working directory is resources/temp/. You \*always\* need to save resources/temp/ at the end of your session!". Below are three buttons: "Save resources/temp/ as ...", "Reset/clear resources/temp/", and "Open previously-saved working directory ...".
- Log output:** A large text area on the right showing the execution log. A red circle highlights the final part of the log, which includes the text "The process finishes at Tue Dec 23 15:44:41 2008" and a series of "Copy resources/..." commands.

# File management: Retrieve the working directory (Step 1)



The screenshot shows the RedundancyMiner application window. The title bar includes the application name "RedundancyMiner". The interface has two tabs: "Default" and "Custom", with "Custom" being the active tab. The main content area is divided into two sections. The top section, titled "Step 1. Choose .CIM file:", contains a button labeled "Click me to select .CIM file" and an empty text field below it. The bottom section, titled "Step 2. Obtain \*genes in categories\* from either:", contains two radio buttons. The first radio button is labeled "same .CIM file chosen in Step 1". The second radio button is labeled ".gce or .txt file" and is selected. Below the radio buttons is a button labeled "Click me to select .gce or .txt file" and an empty text field labeled "The selected \*genes in categories\* file is:". Below these sections is a green "Submit" button. The bottom section, titled "File management", contains the text "Temporary working directory is resources/temp/. You \*always\* need to save resources/temp/ at the end of your session!". Below this text are three buttons: "Save resources/temp/ as ...", "Reset/clear resources/temp/", and "Open previously-saved working directory...". A red arrow points to the "Open previously-saved working directory..." button.

RedundancyMiner

Default Custom

Step 1. Choose .CIM file:

Click me to select .CIM file

Step 2. Obtain \*genes in categories\* from either:

☐ same .CIM file chosen in Step 1

or

☒ .gce or .txt file

Click me to select .gce or .txt file

The selected \*genes in categories\* file is:

Submit

File management

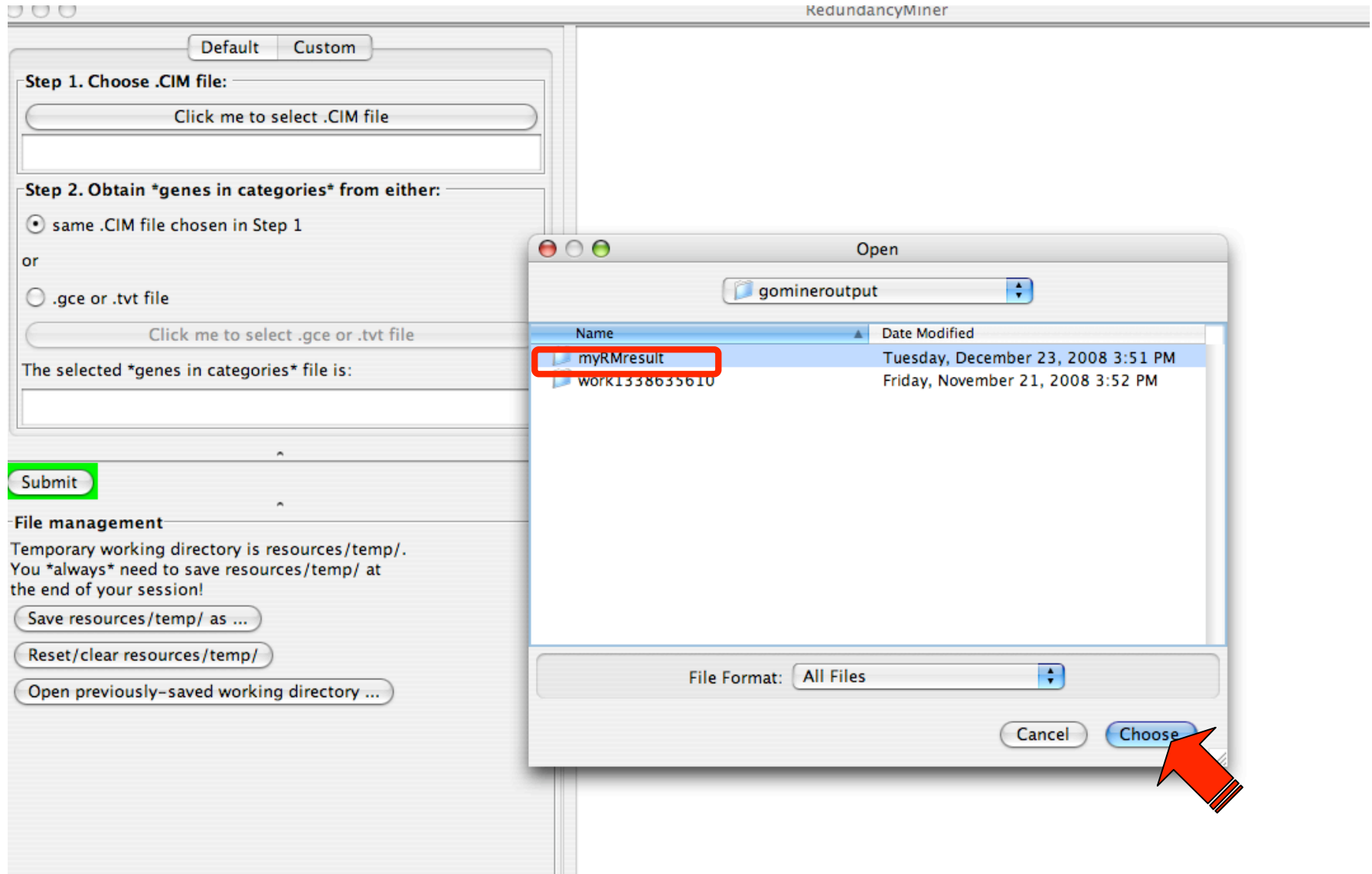
Temporary working directory is resources/temp/.  
You \*always\* need to save resources/temp/ at the end of your session!

Save resources/temp/ as ...

Reset/clear resources/temp/

Open previously-saved working directory...

# File management: Retrieve the working directory (Step 2)



# File management: Retrieval completed

The screenshot displays the RedundancyMiner application interface, which is divided into two main sections: a configuration panel on the left and a terminal window on the right.

**Configuration Panel (Left):**

- Default Custom** tabs are at the top.
- Step 1. Choose .CIM file:** A button labeled "Click me to select .CIM file" is present. Below it, the selected file path is shown: `/All_Genes_wPvals_IntxnEffect.txt.metric.txt.change.series.CIM`.
- Step 2. Obtain \*genes in categories\* from either:** Two radio buttons are shown: "same .CIM file chosen in Step 1" (selected) and ".gce or .txt file". Below them is a button "Click me to select .gce or .txt file". The selected file path is: `35610.dir/All_Genes_wPvals_IntxnEffect.txt.metric.txt.total.txt`.
- Step 3. Set log10(P-Value) Threshold:** A text field shows the threshold: `log10(P-Value) Threshold: -9.223372e+16`.
- Submit** button is highlighted with a green box.
- File management** section: A note states "Temporary working directory is resources/temp/. You \*always\* need to save resources/temp/ at the end of your session!". Below are three buttons: "Save resources/temp/ as ...", "Reset/clear resources/temp/", and "Open previously-saved working directory ...".

**Terminal Window (Right):**

The terminal window, titled "RedundancyMiner", shows a series of commands and their outputs. A red oval highlights the entire terminal area. The commands include copying files to a temporary directory, running a clusterout script, and running a fisher.test script. The output shows the total number of pairs (16653) and a list of genes with their log10(P-Value) thresholds.

```
Copy /Users/hfliu/gomineroutput/myRMresult/All_Genes_wPvals_IntxnEffect.txt.metric.txt.change.series.CIM to
resources/temp/All_Genes_wPvals_IntxnEffect.txt.metric.txt.change.series.CIM
Copy /Users/hfliu/gomineroutput/myRMresult/All_Genes_wPvals_IntxnEffect.txt.metric.txt.change.series.CIM.0.-9.223372E16.0.0.0.0 to
resources/temp/All_Genes_wPvals_IntxnEffect.txt.metric.txt.change.series.CIM.0.-9.223372E16.0.0.0.0
Copy
/Users/hfliu/gomineroutput/myRMresult/All_Genes_wPvals_IntxnEffect.txt.metric.txt.change.series.CIM.0.-9.223372E16.0.0.0.0.INDEX.CIM to
resources/temp/All_Genes_wPvals_IntxnEffect.txt.metric.txt.change.series.CIM.0.-9.223372E16.0.0.0.0.INDEX.CIM
Copy /Users/hfliu/gomineroutput/myRMresult/All_Genes_wPvals_IntxnEffect.txt.metric.txt.total.txt to
resources/temp/All_Genes_wPvals_IntxnEffect.txt.metric.txt.total.txt
Copy
/Users/hfliu/gomineroutput/myRMresult/All_Genes_wPvals_IntxnEffect.txt.metric.txt.total.txt.All_Genes_wPvals_IntxnEffect.txt.metric.txt.chan
ge.series.CIM.0.0.0.-9.223372E16.0.0.0.0.META to
resources/temp/All_Genes_wPvals_IntxnEffect.txt.metric.txt.total.txt.All_Genes_wPvals_IntxnEffect.txt.metric.txt.change.series.CIM.0.0.0.-9.22
3372E16.0.0.0.0.META
Copy
/Users/hfliu/gomineroutput/myRMresult/All_Genes_wPvals_IntxnEffect.txt.metric.txt.total.txt.All_Genes_wPvals_IntxnEffect.txt.metric.txt.chan
ge.series.CIM.0.0.clusterout to
resources/temp/All_Genes_wPvals_IntxnEffect.txt.metric.txt.total.txt.All_Genes_wPvals_IntxnEffect.txt.metric.txt.change.series.CIM.0.0.clustero
ut
Copy
/Users/hfliu/gomineroutput/myRMresult/All_Genes_wPvals_IntxnEffect.txt.metric.txt.total.txt.All_Genes_wPvals_IntxnEffect.txt.metric.txt.chan
ge.series.CIM.0.0.FDRout to
resources/temp/All_Genes_wPvals_IntxnEffect.txt.metric.txt.total.txt.All_Genes_wPvals_IntxnEffect.txt.metric.txt.change.series.CIM.0.0.FDRout
Copy
/Users/hfliu/gomineroutput/myRMresult/All_Genes_wPvals_IntxnEffect.txt.metric.txt.total.txt.All_Genes_wPvals_IntxnEffect.txt.metric.txt.chan
ge.series.CIM.fisher.input.s2s to
resources/temp/All_Genes_wPvals_IntxnEffect.txt.metric.txt.total.txt.All_Genes_wPvals_IntxnEffect.txt.metric.txt.change.series.CIM.fisher.input
.s2s
Copy
/Users/hfliu/gomineroutput/myRMresult/All_Genes_wPvals_IntxnEffect.txt.metric.txt.total.txt.All_Genes_wPvals_IntxnEffect.txt.metric.txt.chan
ge.series.CIM.fisher.output.s2s to
resources/temp/All_Genes_wPvals_IntxnEffect.txt.metric.txt.total.txt.All_Genes_wPvals_IntxnEffect.txt.metric.txt.change.series.CIM.fisher.outp
ut.s2s
Copy /Users/hfliu/gomineroutput/myRMresult/userinputparam.txt to resources/temp/userinputparam.txt
Temporary directory is resources/temp/ ...
Obtaining statistics
The process of obtaining statistics starts at Tue Dec 23 16:14:14 2008

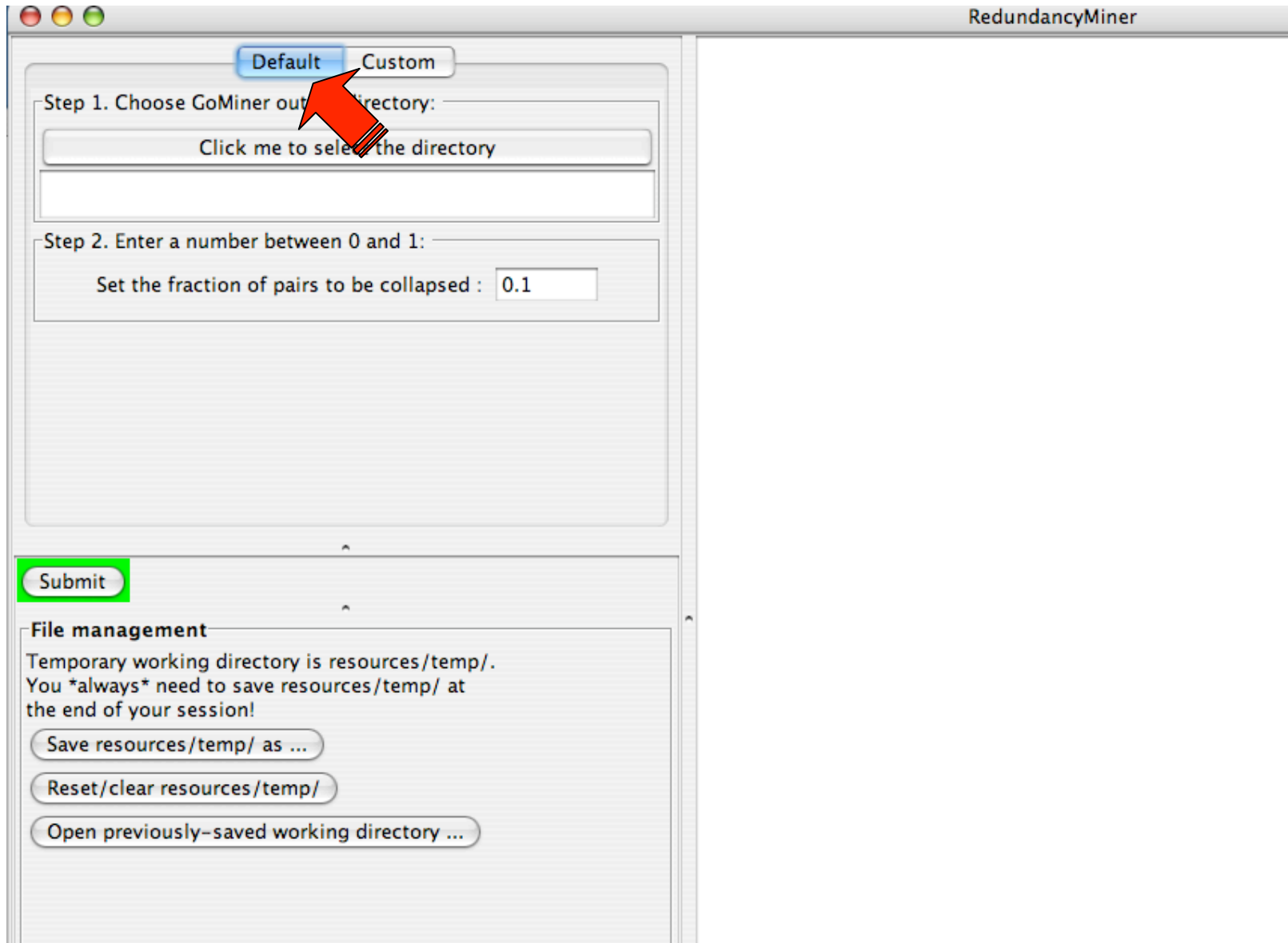
Total number of pairs is 16653
Only show those with P-value less than 0.5
Rank    log10(P-Value)
0       -9.223372e+16
1       -9.223372e+16
3       -9.223372e+16
7       -9.223372e+16
15      -9.223372e+16
```

# Default Mode

The default mode will generate results on the selected HTGM output directory.

- Select HTGM output directory
- Submit the job

# Select Default mode



The screenshot shows the RedundancyMiner application window. At the top, there are two tabs: 'Default' and 'Custom'. The 'Default' tab is selected, and a red arrow points to it. Below the tabs, there are two steps for configuration:

Step 1. Choose GoMiner output directory:

Click me to select the directory

Step 2. Enter a number between 0 and 1:

Set the fraction of pairs to be collapsed :

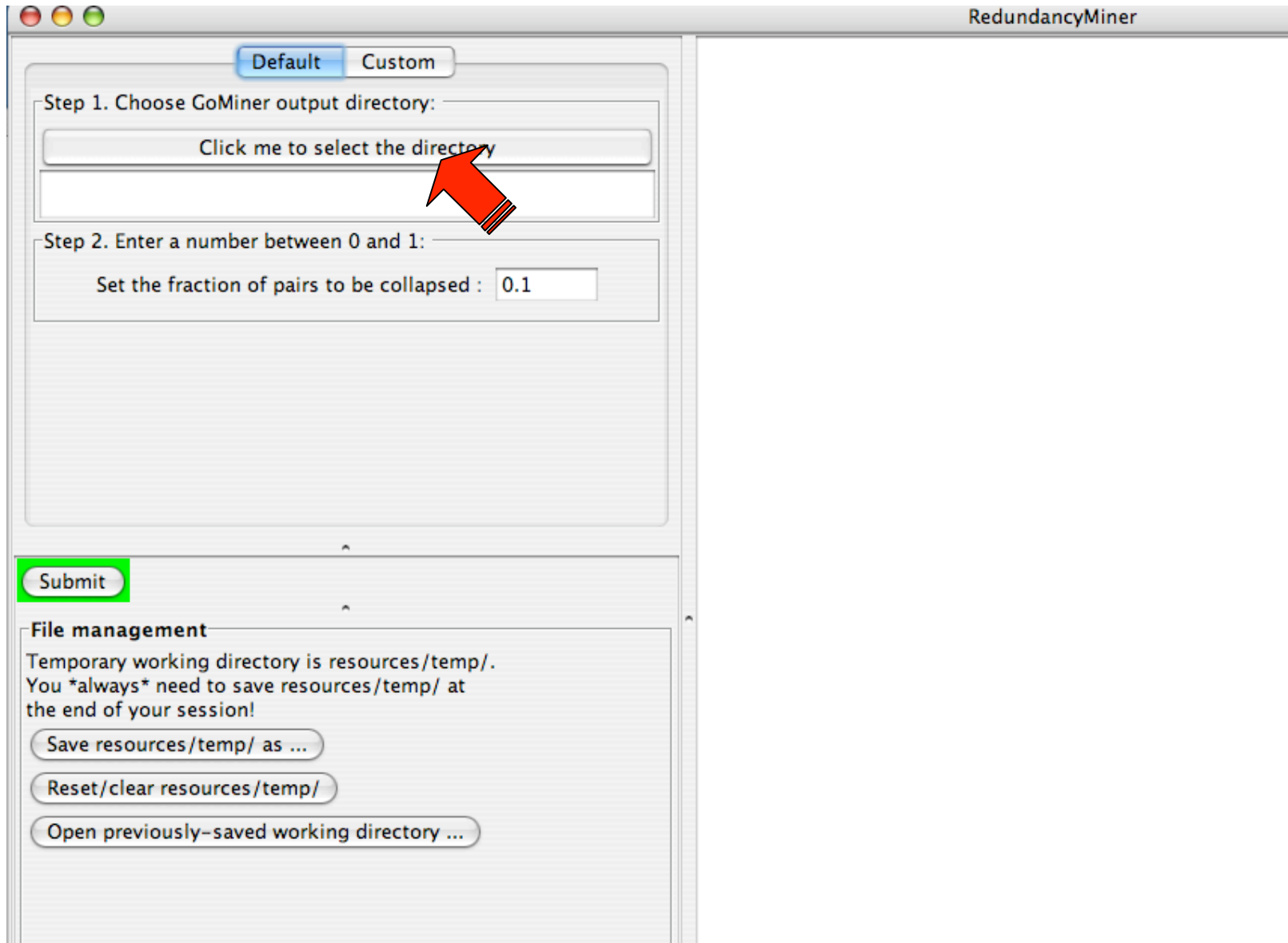
Below these steps, there is a 'Submit' button highlighted with a green border. At the bottom, there is a 'File management' section with the following text:

Temporary working directory is resources/temp/.  
You \*always\* need to save resources/temp/ at the end of your session!

Below this text are three buttons:

- Save resources/temp/ as ...
- Reset/clear resources/temp/
- Open previously-saved working directory ...

# Select HTGM output directory (Step 1)



The screenshot shows the RedundancyMiner application window. At the top, there are two tabs: "Default" (selected) and "Custom". Below the tabs, the interface is divided into two main sections. The top section is for Step 1, "Choose GoMiner output directory:", which contains a button labeled "Click me to select the directory". A red arrow points to this button. The bottom section is for Step 2, "Enter a number between 0 and 1:", which contains a label "Set the fraction of pairs to be collapsed :" and a text input field with the value "0.1". Below these sections is a "Submit" button, which is highlighted with a green border. At the bottom of the window, there is a "File management" section with the text "Temporary working directory is resources/temp/. You \*always\* need to save resources/temp/ at the end of your session!". Below this text are three buttons: "Save resources/temp/ as ...", "Reset/clear resources/temp/", and "Open previously-saved working directory ...".

RedundancyMiner

Default Custom

Step 1. Choose GoMiner output directory:

Click me to select the directory

Step 2. Enter a number between 0 and 1:

Set the fraction of pairs to be collapsed : 0.1

Submit

**File management**

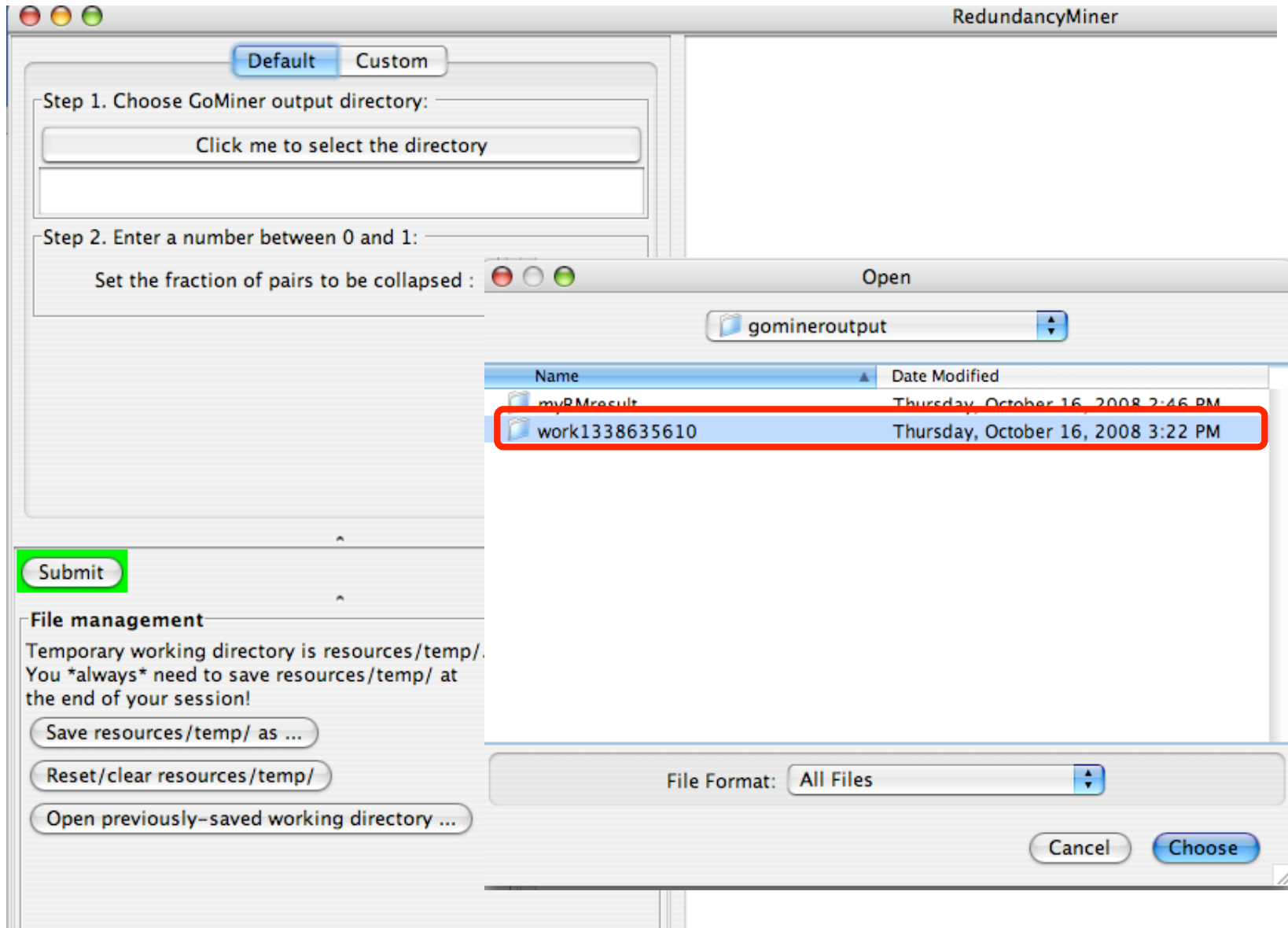
Temporary working directory is resources/temp/.  
You \*always\* need to save resources/temp/ at  
the end of your session!

Save resources/temp/ as ...

Reset/clear resources/temp/

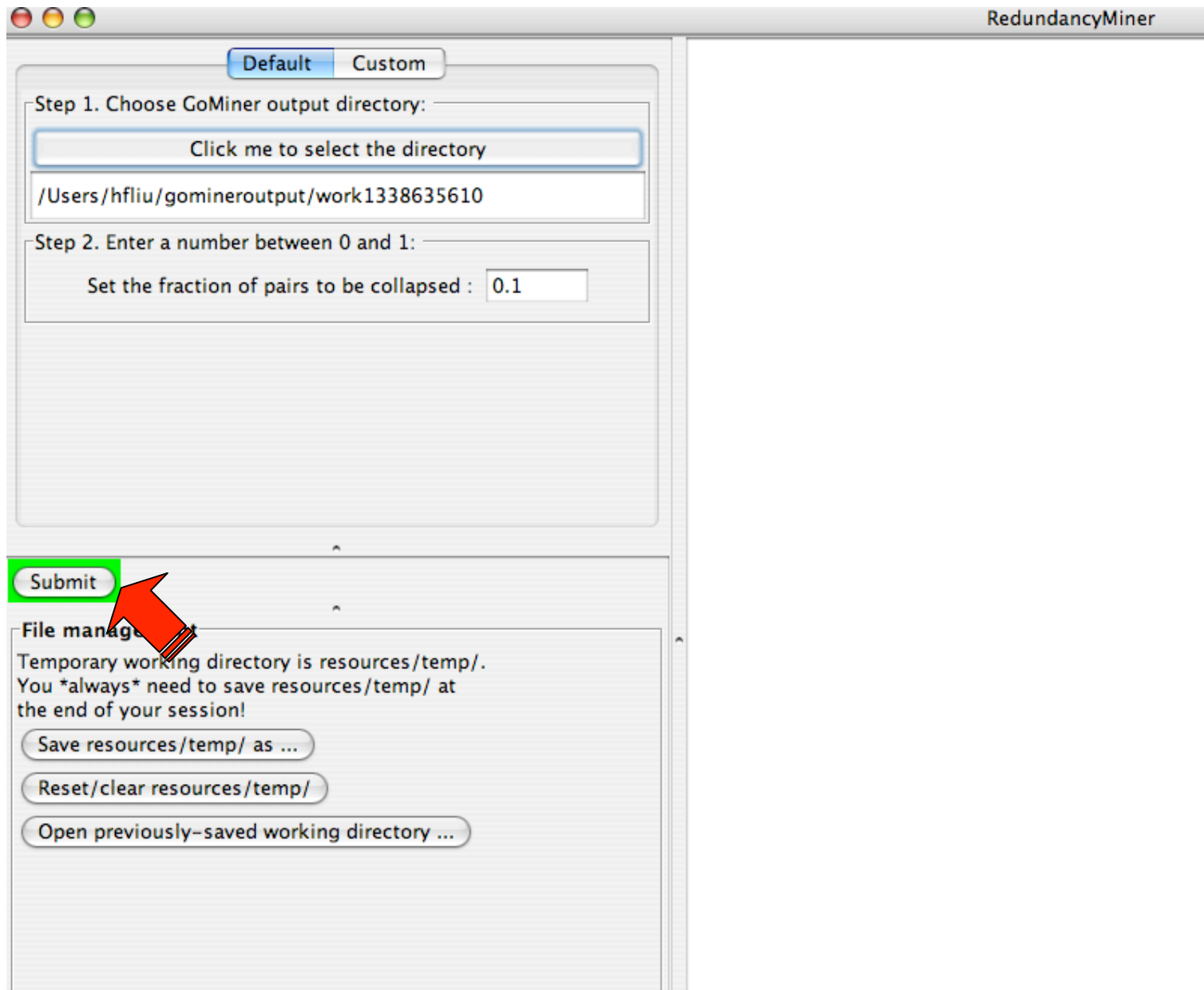
Open previously-saved working directory ...

# Select HTGM output directory (Step 2)





# Submit the job (Step 3)



The screenshot shows the RedundancyMiner application window. The title bar at the top right reads "RedundancyMiner". Below the title bar are two tabs: "Default" (selected) and "Custom".

Step 1. Choose GoMiner output directory:

Step 2. Enter a number between 0 and 1:

Set the fraction of pairs to be collapsed :

**Submit**

**File manager**

Temporary working directory is resources/temp/.  
You \*always\* need to save resources/temp/ at the end of your session!

A red arrow points to the "Submit" button.

# Default mode completed

The screenshot displays the RedundancyMiner application window, which is divided into several sections. At the top, there are two tabs: "Default" (selected) and "Custom".

**Step 1. Choose GoMiner output directory:**

Click me to select the directory

/Users/hfliu/gomineroutput/work1338635610

**Step 2. Enter a number between 0 and 1:**

Set the fraction of pairs to be collapsed : 0.1

**Submit**

**File management**

Temporary working directory is resources/temp/.  
You \*always\* need to save resources/temp/ at the end of your session!

Save resources/temp/ as ...

Reset/clear resources/temp/

Open previously-saved working directory ...

**RedundancyMiner**

Prepare data in /Users/hfliu/gomineroutput/work1338635610...  
Calculate P-Values All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt.change...  
Multiclustering  
The process starts at Fri Nov 21 15:48:57 2008  
Initializing clusters ....  
Building clusters ....  
Printing clusters ....  
The process finishes at Fri Nov 21 15:49:07 2008  
Collapsing CIM according to clustering results and storing in their corresponding directory...  
The process starts at Fri Nov 21 15:49:07 2008  
Processing  
/Users/hfliu/gomineroutput/work1338635610/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt1338635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metri  
c.txt.change.series.CIM at Fri Nov 21 15:49:07 2008  
The number of categories:  
Before collapsing: 183  
After collapsing: 162  
Compression ratio: 1.13  
Processing  
/Users/hfliu/gomineroutput/work1338635610/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt1338635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metri  
c~\_0000000001.txt.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric~\_0000000001.txt.change.gce.CIM at Fri Nov 21 15:49:26 2008  
Processing  
/Users/hfliu/gomineroutput/work1338635610/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt1338635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metri  
c~\_0000000002.txt.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric~\_0000000002.txt.change.gce.CIM at Fri Nov 21 15:49:28 2008  
Processing  
/Users/hfliu/gomineroutput/work1338635610/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt1338635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metri  
c~\_0000000003.txt.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric~\_0000000003.txt.change.gce.CIM at Fri Nov 21 15:49:34 2008  
Processing  
/Users/hfliu/gomineroutput/work1338635610/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt1338635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metri  
c~\_0000000004.txt.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric~\_0000000004.txt.change.gce.CIM at Fri Nov 21 15:49:44 2008  
Processing  
/Users/hfliu/gomineroutput/work1338635610/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt1338635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metri  
c~\_0000000005.txt.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric~\_0000000005.txt.change.gce.CIM at Fri Nov 21 15:49:55 2008  
Processing  
/Users/hfliu/gomineroutput/work1338635610/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt1338635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metri  
c~\_0000000006.txt.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric~\_0000000006.txt.change.gce.CIM at Fri Nov 21 15:50:09 2008  
Processing  
/Users/hfliu/gomineroutput/work1338635610/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt1338635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metri  
c~\_0000000007.txt.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric~\_0000000007.txt.change.gce.CIM at Fri Nov 21 15:50:25 2008  
Processing  
/Users/hfliu/gomineroutput/work1338635610/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt1338635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metri  
c~\_0000000008.txt.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric~\_0000000008.txt.change.gce.CIM at Fri Nov 21 15:50:48 2008  
Processing  
/Users/hfliu/gomineroutput/work1338635610/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt1338635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metri  
c~\_0000000009.txt.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric~\_0000000009.txt.change.gce.CIM at Fri Nov 21 15:51:09 2008  
Processing  
/Users/hfliu/gomineroutput/work1338635610/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt1338635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metri  
c~\_0000000010.txt.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metric~\_0000000010.txt.change.gce.CIM at Fri Nov 21 15:51:32 2008  
Processing  
/Users/hfliu/gomineroutput/work1338635610/All\_Genes\_wPvals\_IntxnEffect.txt.metric.txt1338635610.dir/All\_Genes\_wPvals\_IntxnEffect.txt.metri